# Opti{c,m}al: Optical/Optimal Routing in Massively Dense Wireless Networks

Roberto Catanuto, Stavros Toumpis, Giacomo Morabito

*Abstract*— We study routing for massively dense wireless networks, i.e., wireless networks that contain so many nodes that, in addition to their usual *microscopic* description, a novel *macroscopic* description becomes possible. The macroscopic description is not detailed, but nevertheless contains enough information to permit a meaningful study and performance optimization of the network. Within this context, we continue and significantly expand previous work on the analogy between optimal routing and the propagation of light according to the laws of Geometrical Optics. Firstly, we pose the analogy in a more general framework than previously, notably showing how the eikonal equation, which is the central equation of Geometrical Optics, also appears in the networking context. Secondly, we develop a methodology for calculating the cost function, which is the function describing the network at the macroscopic level. We apply this methodology for two important types of networks: bandwidth limited and energy limited.

*Index Terms*—Massively dense networks, Optics, Routing, Trajectory Based Forwarding, Wireless networks.

## I. INTRODUCTION

### A. Trajectory Based Forwarding and Massively Dense Networks

In most routing protocols for wireless networks, each route is described in terms of the sequence of nodes that a packet must follow from its source to its destination. The node sequence is either determined at the source (for source driven protocols) or along the way to the destination (for table driven protocols) [1]. However, in both cases, the network often suffers from the *scalability problem*: As the number of nodes in the network increases, the number of nodes participating in a typical route also increases, and so the route becomes more susceptible to node mobility and channel fluctuations. Therefore, more and more bandwidth must be used to fix errors and keep the routes updated [2].

In order to solve the scalability problem, the authors of [3] have endeavored to turn the large number of nodes into an advantage. Notably, they consider a *dense* network, and proceed to propose routing by means of *Trajectory Based Forwarding* (TBF). The network is dense in the sense that it consists of a very large number of nodes so that each node typically has within its communication range several neighboring nodes, that it can use to forward packets toward all directions. In this setting, the route is specified not in terms of a list of nodes, to be visited consecutively on the way to the destination, but rather as a continuous curve, called the *trajectory*, that starts at the physical location of the source and ends at the physical location of the destination. When the packet arrives at an intermediate node, that node forwards the packet to a node that happens, at that time, to be

further ahead along the trajectory of the packet, according to a greedy forwarding rule. As continuous curves can be described compactly, and there is no need to specify beforehand the nodes that participate in the route or maintain routing tables, TBF scales with the network size much better than traditional routing protocols. TBF may be thought of as a generalization of geographic routing [4], with which the trajectory is the straight line connecting the source and the destination.

The concept of dense networks took a significant evolutionary step forward with the introduction of *massively dense* networks [5], [6], [7], [8], [9], [10]. These networks consist of such a large number of nodes, that we can take a novel *macroscopic* view (i.e., bird's eye view) of the network, in addition to the more traditional *microscopic* view, and approximate the various microscopic quantities of interest with concise macroscopic equivalents.

The precise microscopic and macroscopic quantities considered depend on the particular problem at hand. Typical examples of microscopic quantities are node positions, data transmission rates at individual relaying nodes, and the data creation rates at individual source nodes. Their corresponding macroscopic quantities are the density of nodes (a scalar measured in nodes per square meter), the flow of traffic (a vector measured in bps per meter), and the density of data creation at the sources (measured in bps per square meter), all functions of the location $(x, y)$.

Focusing on the macroscopic and ignoring the microscopic quantities involves losses in describing the network. For example, if we describe the network deployment in terms of a node density function, we no longer keep track of the particular placements of individual nodes. On the other hand, the macroscopic quantities retain sufficient detail so as to capture the fundamental properties of the network, and permit the derivation of meaningful and insightful results that describe adequately the behavior of the network.

### B. Routing by Analogy to Geometrical Optics

Of particular interest to our work is the line of investigation initiated in [5], and in which the term 'massively dense' first appeared. There, the objective is to find the route between a source and a destination that requires the minimum number of hops. The network topology is described in terms of the spatial node density function. It is assumed that communication is restricted between nearest neighbors, so that in areas of higher node density each hop corresponds to a smaller physical distance. As a result, packets should try to steer away from high node density areas. The author shows that, in the massively

dense regime, the optimal trajectory satisfies the same set of equations that are satisfied in Geometrical Optics by a ray of light that travels from the source to the destination, assuming the massively dense network is substituted by a properly specified inhomogeneous optical medium. Therefore, we can use the theory, tools, and intuition that exist in Geometrical Optics for solving a problem in networking.

Despite the high degree of originality of [5], the scope of that work is limited: The aim is to find the trajectory from the source to the destination that uses the minimum number of hops, assuming communication is restricted between nearest neighbors. However, in wireless networks communication is typically not limited between nearest neighbors, but rather between nodes that are close enough to hear each other's transmissions. Furthermore, minimizing the number of hops of a route is only one possible objective, very frequently not the most relevant: other objectives are the minimization of delay, bandwidth, energy, etc.

These limitations motivated the work in [11]. There, it is assumed that each transmission has a fixed energy cost, and nodes can communicate with each other as long as they are within some maximum distance from each other. The aim is to find the route that uses the least amount of energy. The authors show that this setting actually favors the routing of packets through areas of high node density. Again, they show that in the massively dense regime, the optimal routes satisfy the same set of equations as those satisfied by a ray of light traveling through a properly defined inhomogeneous optical medium. In addition, they place the problem in the context of TBF: each ray of light is a trajectory that the packets must follow according to a trajectory based forwarding rule. (The details of how forwarding is performed on the microscopic level were ignored in [5].)

Here, we continue on the line of investigation initiated in [5] and continued in [11], making a number of contributions in two main directions. Firstly, we pose the analogy between Geometrical Optics and massively dense networks in a more general framework, that permits its use under a wider variety of problems. Notably, we specify the networking equivalent of the eikonal equation of Optics. Our framework allows to identify the optimal trajectory in two relevant contexts: a simple routing problem, in which information must be transferred by a given source to a given destination, and a distributed routing problem, in which the source or the destination can be any from a set of known nodes. Observe that the second problem cannot be handled with the frameworks considered in [5] and [11].

Secondly, we develop a methodology for calculating the cost function, which is the key macroscopic quantity for our work, in several scenarios, and we exhibit this methodology in two cases of high relevance: a bandwidth limited network model, and an energy limited network model.

The rest of this paper is organized as follows. To keep the text self-contained, in Section II we briefly review the basic concepts of Geometrical Optics that we use. In Section III we introduce a macroscopic model of the network which
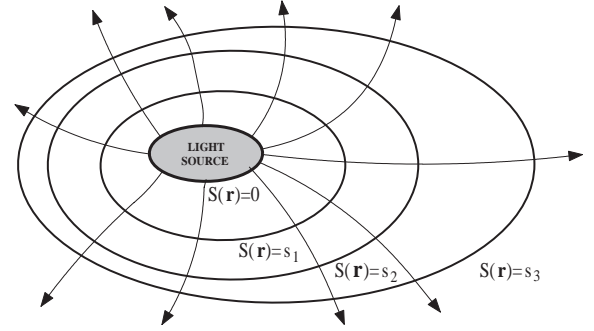


Fig. 1. Wavefronts and rays emanating from a light source.

formalizes the analogy between Geometrical Optics and the problem of routing in massively dense wireless networks. The macroscopic model needs the specification of an appropriate cost function to be derived using a microscopic model of the network, and so in Section IV we derive the cost function for a bandwidth limited model and an energy limited model. The proposed analogy is assessed in Section V where simulation results show that the trajectories derived through the proposed analogy closely follow the route obtained by applying shortest path routing algorithms such as the Bellman-Ford algorithm. Finally, open research issues are given in Section VI.

## II. BACKGROUND ON GEOMETRICAL OPTICS

In this section we provide a brief overview of Geometrical Optics [12], [13], and in particular those aspects that are related to our work. As our networks are 2-dimensional, we discuss Optics in a 2-dimensional setting.

Geometrical Optics describes the propagation of electromagnetic fields when the wavelength is much smaller than the dimensions of the features of the environment or, more technically, in the limit where its wavelength goes to zero. In this approximating case, the laws of propagation can be formulated in the language of geometry, hence the name of this theory. Note that the wavelengths of visible light are on the order of micrometers, and therefore visible light is very well described by Geometrical Optics.

The emanation of light from a light source is described in terms of the *eikonal function* $S(\mathbf{r})$ which is measured in meters and equals the time it takes for the light to go from the source to point $\mathbf{r}$, multiplied with the speed of light in free space $c$. Therefore, the eikonal function $S(\mathbf{r})$ represents the distance that light will cover *in free space*, at the time it takes light to travel, *through the medium*, from the source to the point $\mathbf{r}$. Surfaces of the form $S(\mathbf{r}) = k > 0$ are called *wavefronts*, and represent loci of points at which light, leaving the source at time 0, will arrive at time $\frac{k}{c}$.

It follows from Maxwell's equations that the function $S(\mathbf{r})$ satisfies the following non-linear partial differential equation, known as the *eikonal equation*:

$$(\frac{\partial S}{\partial x})^2 + (\frac{\partial S}{\partial y})^2 = n^2 \Leftrightarrow (\nabla S)^2 = n^2 \Leftrightarrow |\nabla S| = n. \quad (1)$$

In the above, $n = n(\mathbf{r}) \geq 1$ is the *refractive index*, a unit-less scalar quantity that captures the properties of the medium at location $\mathbf{r}$. The vector $\nabla S \triangleq \frac{\partial S}{\partial x}\hat{\mathbf{x}} + \frac{\partial S}{\partial y}\hat{\mathbf{y}}$ is the gradient of $S$. Its direction gives the direction along which $S$ increases the fastest, and its absolute value $|\nabla S|$ gives the rate of this increase. To understand the meaning of (1), note that at a point $\mathbf{r}$, $|\nabla S| = \frac{dS}{ds} = \frac{dt \times c}{ds}$, where $ds$ is an incremental distance along the direction of maximum increase of $S$, $dS$ is its increase along $ds$, and $dt$ is the time it takes the light to cover distance $ds$. Applying (1), it follows that $\frac{ds}{dt} = \frac{c}{n(\mathbf{r})}$, i.e., equation (1) specifies that at the location $\mathbf{r}$ light travels with speed $\frac{c}{n(\mathbf{r})}$. The initial condition needed for solving (1) is that $S(\mathbf{r}) = 0$ along the set of points $\Omega$ representing the source. Having this initial condition, the eikonal equation can then be solved for all points $\mathbf{r}$ outside the source.

We briefly note that the eikonal equation is of interest not only in the context of Geometrical Optics, but can be used to model a large number of disparate phenomena in disciplines such as seismology, fluid mechanics, and material science [14]. As an off-the-wall example, the eikonal equation can be used to predict the way a large group of termites will eat up wood, assuming that they are initially placed in a domain $\Omega$, and are capable of eating up the wood at point $\mathbf{r}$ with speed $\frac{c}{n(\mathbf{r})}$, where $c$ is a nominal speed, and $n(\mathbf{r}) > 0$ is the wood hardness.

Having defined the eikonal function and the related notion of wavefronts, we can define *light rays* as the set of curves that originate at the light source and move away from it, while keeping constantly orthogonal to wavefronts. An example is shown in Fig. 1. Intuitively, light rays are the paths that light takes as it emanates away from the source.

By applying the eikonal equation (1), it can easily be seen [12] that light rays satisfy the following non-linear second-order differential equation:

$$\frac{d}{ds}\left(n\frac{d\mathbf{r}}{ds}\right) = \nabla n, \tag{2}$$

where $s$ is the length of the ray measured from a fixed point to position $\mathbf{r}$, $\frac{d\mathbf{r}}{ds}$ is the direction of the ray at $\mathbf{r}$, $n(\mathbf{r})$ is the refractive index at $\mathbf{r}$, and $\nabla n = \frac{\partial n}{\partial x}\hat{\mathbf{x}} + \frac{\partial n}{\partial y}\hat{\mathbf{y}}$ is its gradient. Note that the gradient $\nabla n$ points toward the direction where the refractive index increases. Therefore, (2) expresses the fact that light rays tend to bend toward the direction where the medium is optically denser, i.e., the index of refraction has a higher value. Equation (2) holds wherever the refractive index is a continuous function of the position. Locations with discontinuities can be handled by applying (1) together with limiting arguments, leading to Snell's laws of reflection and refraction [12].

Next, let the *optical length* $[AB]_C$ of a curve $C$ connecting two points $A$ and $B$ be the line integral

$$[AB]_C = \int_A^B n(\mathbf{r})\, ds \tag{3}$$

along the curve $C$. Since the speed with which light moves at a point $\mathbf{r}$ is $u = \frac{c}{n(\mathbf{r})}$, the time needed by light to go from $A$
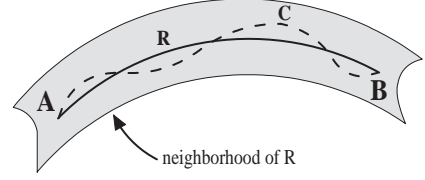


Fig. 2. Illustration of the principle of Fermat: Light traveling from $A$ to $B$ along any small perturbation $C$ of a light ray $R$ would necessarily take more time.

to $B$ or vice versa along the curve $C$ equals

$$\int_A^B dt = \int_A^B \frac{1}{u}\, ds = \frac{1}{c}\int_A^B n(\mathbf{r})\, ds = \frac{[AB]_C}{c}.$$

Therefore, the optical length of a curve equals the time it would take light to travel across the curve $C$, multiplied by the speed of light in free space $c$.

Now let the source of light be at a point $A$, and let $R$ be a curve starting at $A$ and ending at some other point $B$, as shown in Fig. 2. The *principle of Fermat* states that the curve $R$ will be a ray of light if and only if its optical length is smaller than the optical length of any other curve $C$ that passes through $A$ and $B$ and exists within a sufficiently small neighborhood of the ray $R$. In other words, if we introduce a sufficiently small perturbation in a ray $R$, resulting in a new curve $C$, then necessarily the new curve will have a larger optical length. As the optical length is proportional to the time needed to go from $A$ to $B$, the principle of Fermat asserts that light takes the fastest route.

## III. MACROSCOPIC NETWORK MODEL

Our network consists of nodes placed over a 2-dimensional domain $\mathcal{D}$. We assume that there are so many nodes, that the typical distance between nearest neighbors is incremental with respect to the dimensions of $\mathcal{D}$.

We associate with the movement of packets over the network a consumption of resources which is cumulative with the distance over which the packet is transported. The resources could be the bandwidth used, the energy wasted, the time needed for the transport, or some other composite of the above. However, in this section, we refrain from further specifying the nature of the resources, and we keep the discussion abstract. We model the use of resources in terms of a scalar *cost function* $c(\mathbf{r})$, which is unit-less, and is defined as follows: Let $\epsilon$ be the incremental length of a small line segment centered at $\mathbf{r}$. Let $dc(\epsilon, \mathbf{r})$ be the incremental cost associated with transporting a packet along this incremental line segment. Let $c_{\text{nominal}} \times \epsilon$ be the cost incured by transporting the packet in some given *nominal network* by a distance $\epsilon$. The cost function is defined as follows:

$$c(\mathbf{r}) = \lim_{\epsilon \to 0} \frac{dc(\epsilon, \mathbf{r})}{c_{\text{nominal}} \times \epsilon}. \tag{4}$$

For example, a cost function $c(\mathbf{r}) = 2$ means that it is twice as costly to move packets along the location $\mathbf{r}$ of our network with respect to the nominal network.

Initially, this definition of the cost function may seem counterintuitive. Indeed, it would make more sense to simply define the cost function as the incremental cost over the incremental distance, without dividing by $c_{\mathrm{nominal}}$. The advantages of our approach are two-fold: firstly, in this manner the cost function is unit-less, irrespective of the precise nature of the cost. This makes the definition more general. Secondly, we will soon show an analogy of our setting with Optics, where the nominal network is mapped to free space, and the cost function is mapped to the refractive index. As the refractive index is unit-less, making the cost function also unit-less makes the analogy tighter.

Using the definition of the cost function (4), it follows that the cost to move a packet from point $A$ to point $B$, along a curve $C$ of non-incremental length, is given by the line integral of the cost function along $C$:

$$[AB]_C = \int_A^B c(\mathbf{r})\,d\mathbf{r}. \qquad (5)$$

The cost is measured in meters, and represents the distance that will be covered in the nominal network, using the same volume of resources.

We apply the above concepts in two relevant scenarios:

- **Simple Routing Problem:** Let $A$ be a node in the network. Given any other node $B$ in the network, what is the minimum cost route connecting the two, and what is the associated minimum cost?
- **Distributed Routing Problem:** Let $\Omega \subset \mathcal{D}$ be a subdomain of $\mathcal{D}$ containing a set of nodes in the network. Given any other node $B$ in the network, what is the minimum cost route connecting $B$ with $\Omega$, and what is the associated minimum cost?

Note that the second question is fundamentally different from the first question, because in the first case we want to find a route connecting two known locations, but in the second case the one end of the route is not known. The second case is relevant where we need to find a minimum cost route, but we have some freedom in selecting one of its end points. For example, the set $\Omega$ might be a collection of points corresponding to the cluster heads of a wireless ad hoc network. In such a case, the rest of the nodes will want to communicate with that of the cluster heads which incurs the minimum cost. As another example, the set $\Omega$ might represent a distributed set of sinks that are placed in a wireless sensor network. Individual sensors will want to relay the information they have collected to the nearest sink.

### A. Analogy with Geometrical Optics

Both problems involve the minimization of the integral (5) over all possible curves $C$, subject to constraints on the start and end of the curves, and can be solved with Calculus of Variations techniques. Luckily for us, both of these two problems have been solved and extensively investigated in the context of Optics.

In more detail, optical media are described in terms of the refractive index $n(\mathbf{r})$: in the time it takes light to cover a distance of $\epsilon$ in the medium, light can cover a distance of $\epsilon n(\mathbf{r})$ in free space. Networks, on the other hand, are described in terms of the cost function: with the cost it takes the network to transport the packet across a distance $\epsilon$, the packet could be transported across a distance $\epsilon c(\mathbf{r})$ in the nominal network. Therefore, going from Optics to Networking, the optical medium becomes the network, the refractive index becomes the cost function, the optical length of a curve, given by (3), becomes the cost to transport a packet along that curve, given by (5), and free space becomes the nominal network.

We can then solve the Simple Routing Problem by framing it into the context of Optics: Finding the curve with the minimum cost is equivalent to finding the curve with the minimum optical length in an optical medium with refractive index $n(\mathbf{r}) = c(\mathbf{r})$ and a source of light at point $A$. However, as already discussed, Fermat's principle specifies that this curve is actually a ray of light that starts from $A$ and goes to $B$! Therefore, the optimal route satisfies the differential equation (2), if we substitute $n(\mathbf{r})$ with $c(\mathbf{r})$, together with the initial conditions that it must start at $A$ and end at $B$.

We can also solve the Distributed Routing Problem in a similar manner. In particular, we need the minimum costs, and the routes that achieve them, between the region $\Omega$ and all other points in the network. In the context of Optics, we need to find the minimum optical lengths, and their corresponding rays, between a distributed light source occupying $\Omega$ and all other points $B$ in an optical medium whose refractive index is $n(\mathbf{r}) = c(\mathbf{r})$. By definition, these minimum optical lengths are given by the eikonal function, which in turn can be found by solving the eikonal equation (1). Knowing the eikonal function, we can costruct the wavefronts and from these the rays that connect the source with any part of the network.

## IV. MICROSCOPIC NETWORK MODELS

In the previous section we developed a macroscopic network model, built around the notion of an abstract cost function, which we left unspecified. In this section, we develop two microscopic models, specifying, among others, the resources consumed by each transmission, the model for node placement, and the rules of forwarding packets. We then calculate the cost function that corresponds to each of these microscopic models.

### A. Bandwidth Limited Networks

We consider first the case in which the most scarce resource is bandwidth and therefore the objective of the routing protocol is to minimize bandwidth consumption.

**Node Deployment:** We assume that the nodes are placed randomly, according to a spatial Poisson distribution [15], [16] with density $\lambda(\mathbf{r})$ measured in nodes per square meter and defined by the following limit:

$$\lambda(\mathbf{r}) = \lim_{\epsilon \to 0} \frac{E[N(\mathcal{B}(\mathbf{r}, \epsilon))]}{\pi \epsilon^2},$$

where $\mathcal{B}(\mathbf{r}, \epsilon)$ is the ball centered at $\mathbf{r}$ with radius $\epsilon$, $N(\mathcal{B}(\mathbf{r}, \epsilon))$ the number of nodes within that ball, and $E[\cdot]$ denotes the expectation of a random variable.
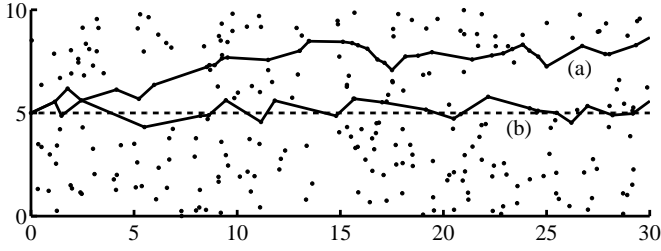
Fig. 3. Packet forwarding along a straight trajectory according to two different forwarding rules: **Rule A:** (route (a)) The cost-progress ratio $\frac{c_t}{x}$ must be minimized. **Rule B:** (route (b)) The cost-progress ratio $\frac{c_t}{x}$ must be minimized, but packets must stay also close to the trajectory.

**Transmission cost:** We assume that the nodes communicate with each other through a common wireless channel of finite bandwidth. When a node $T$ transmits to a node $R$, nodes in the vicinity must refrain from transmitting, in order to avoid collisions.

In this setting, it is reasonable to assume that each transmission over a distance $d$ occupies an area proportional to $d^2$. An intuitive justification for this is as follows: If a node $T$ transmits to a node $R$ over some distance $d$, then it is necessary for all other potential transmitters that are in a neighborhood of the receiver to refrain from transmitting. Intuitively, the radius $r$ of this neighborhood is proportional to the distance $d$. Indeed, the larger $d$ is, the larger $r$ should be to ensure a minimum acceptable Signal to Interference Ratio (SIR) at the receiver. Therefore, the area of the neighborhood is $kd^2$, where $k$ is a unit-less constant that depends on the sensitivity of the receiver, the rate used, etc. A rigorous derivation of the consumed area for a particular reception model exists in [17], but is not replicated here for lack of space.

Clearly, as the areas occupied by communicating pairs in a network increase in size, the number of simultaneous transmissions decreases, as the total occupied area can not exceed the total area of the network. In turn, this drives the aggregate throughput down[1]. It would be reasonable for network designers to keep the size of the occupied areas as small as possible, in order to maximize throughput, and so it is reasonable to define the *transmission cost* of a transmission over a distance $d$ to be

$$c_t = d^2. \tag{6}$$

**Forwarding Rule:** To complete our microscopic model we must also specify the rule according to which nodes forward packets toward their destination. Microscopic parts of trajectories can be viewed as straight lines, therefore it suffices to have a rule for forwarding packets along a given direction. With no loss of generality, let this direction be that of the positive $x$-axis.

Given that we have associated each transmission over a distance $d$ with a cost $c_t = d^2$, a reasonable first choice for

---

[1]We note that the proof for the upper bound of the capacity of a wireless network in [17] was based on exactly this idea.

the node is to forward the packet to that of its neighbors closer to the destination (along the trajectory) for which the ratio

$$r \triangleq \frac{c_t}{x_2 - x_1} = \frac{d^2}{x_2 - x_1}$$

is minimized. In the above, $d$ is the distance between the node and its neighbor, $x_1$ is the $x$-coordinate of the node, and $x_2$ is the $x$-coordinate of the neighbor. The difference $x_2 - x_1 > 0$ represents the progress toward the destination at the end of the trajectory. Therefore, we would like to strike a balance between incuring a small transmission cost, but also achieving a substantial progress toward the destination. We will refer to this forwarding rule as Rule A. In Fig. 3 we plot a sample route (route (a)) that would result from this forwarding rule.

As shown in the figure, a problem with Rule A is that it may cause the packet to deviate significantly from its trajectory, over the course of many hops. (In fact, it is intuitively clear that the deviation from the trajectory can be modeled by a random walk where the steps take continuous values.) Therefore, we are motivated to consider the following modification: if the node forwarding the packet is on the left (right) of the trajectory and at some distance $v$ from it, then the next node to receive the packet can not be also on the left (right) of the trajectory and at some distance $v' > v$ from it. The rest of the nodes are acceptable as next hops, and the one minimizing the cost-progress ratio should be chosen. An example route appears in Fig. 3 (route (b)). As the figure shows, the new rule leads to the packets following the trajectory much more closely.

A disadvantage of Rule B with respect to Rule A is that we are excluding from our consideration neighbors with a very small associated cost-progress ratio, but who happen to increase the deviation from the ideal trajectory even very modestly. The observant reader will notice a few such examples in Fig. 3. Obviously, there is significant room for improving both Rule A and Rule B. As the question of choosing the next hop within the context of TBF and geographic routing has already attracted significant research interest (see for example [18] and the references therein), we do not pursue this issue further. In the following, we will assume that Rule B is used.

**Forwarding Rule Performance:** Next, we evaluate the performance of Rule B in terms of the expected progress and the expected cost per hop. With no loss of generality, we assume that the node having the packet is placed at the origin, and that the trajectory is along the positive $x$-axis, and described by the equation $y = \text{const} > 0$. The progress is then simply the $x$-coordinate of the next node, and in order to have a positive progress toward the destination, the next hop must have a positive $x$-coordinate. In addition, in order to satisfy Rule B, it must also have a positive $y$-coordinate. Therefore, only nodes in the positive quadrant will be considered for forwarding the packet.

As the placement of nodes is random, the location of the next node will also be random, and so its associated progress, $y$-coordinate, distance from the origin, cost, and cost-progress ratio are random variables, which we denote by $X$, $Y$, $D$, $C$,
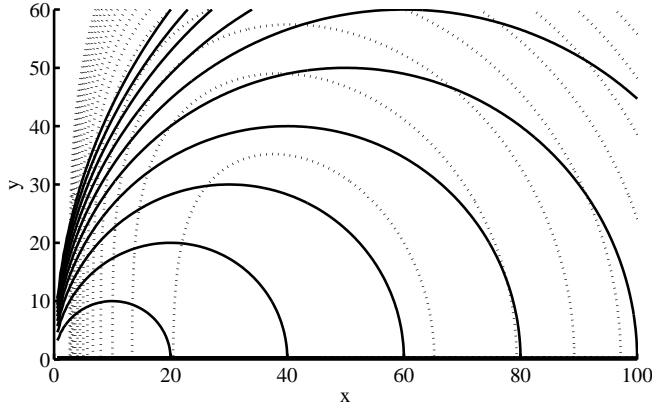
Fig. 4. Loci of constant cost-progress ratio in the bandwidth limited case (continuous lines) and energy-limited case (dashed lines).

and $R$ respectively.

The first step is to find the loci of constant cost-progress ratio. With straightforward algebra, it follows that:

$$\frac{d^2}{x} = r \Leftrightarrow (x - \frac{r}{2})^2 + y^2 = (\frac{r}{2})^2.$$

Therefore, as shown in Fig. 4, the loci of constant ratio are semi-circles centered at point $(\frac{r}{2}, 0)$ with radius $\frac{r}{2}$. Note that as $r \to 0$, the corresponding locus implodes toward the origin. Therefore, to minimize the cost-progress ratio, the node would like to avoid transmitting to nodes further away. This is consistent with the long-known fact that many short hops are better than a few long ones, if the bandwidth is limited [17], [19].

Knowing the shape of the loci, we can calculate the distribution of the ratio $R$. The ratio will be larger than $r$ if and only if there is no node within a semi-circle of radius $\frac{r}{2}$. As the nodes follow a Poisson spatial distribution with density $\lambda(\mathbf{r})$, this happens with probability $\exp[-\frac{\lambda(\mathbf{r})\pi r^2}{8}]$ . Therefore,

$$F_R(r) = 1 - e^{-\frac{\lambda(\mathbf{r})\pi r^2}{8}} \Leftrightarrow f_R(r) = \frac{r\lambda(\mathbf{r})\pi}{4} e^{-\frac{\lambda(\mathbf{r})\pi r^2}{8}}.$$

So the ratio $R$ follows the Rayleigh distribution, with expectation $E[R] = \sqrt{\frac{2}{\lambda(\mathbf{r})}}$.

Next, we calculate the expectation $E[X]$ of the progress, using conditioning on $R$. In particular, let us assume that $R = r$. It then follows that the next node is on the circumference of the semi-circle centered at $(\frac{r}{2}, 0)$ and with radius $\frac{r}{2}$. By symmetry, it follows that the conditional expectation of its $x$-coordinate will be $E[X|R = r] = \frac{r}{2}$. Therefore:

$$
\begin{aligned}
E[X] &= \int_0^\infty E[X|R = r]f_R(r)\,dr \\
&= \int_0^\infty \frac{r^2\lambda(\mathbf{r})\pi}{8} \exp[-\frac{\lambda(\mathbf{r})\pi r^2}{8}]\,dr = \frac{1}{\sqrt{2\lambda(\mathbf{r})}}.
\end{aligned}
$$

Finally, we calculate the expectation $E[C] = E[D^2]$ of the cost, again by conditioning on the value of the ratio $R$. For this, we need to calculate $E[D^2|R = r]$. We note that
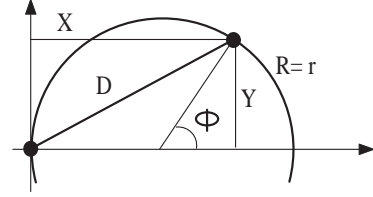


Fig. 5. Random variables associated with the next hop.

$D^2 = \frac{R^2}{2}(1 + \cos\Phi)$ where, as shown in Fig. 5, $\Phi$ is the angle formed by the positive $x$-axis and the line interval connecting the points $(X, Y)$ and $(\frac{r}{2}, 0)$. Given that $R = r$, by symmetry considerations the angle $\Phi$ is uniformly distributed in $(0, \pi)$. It follows that

$$E[D^2|R = r] = \frac{1}{\pi}\int_0^\pi \frac{r^2}{2}(1 + \cos\phi)\,d\phi = \frac{r^2}{2}.$$

Therefore,

$$
\begin{aligned}
E[D^2] &= \int_0^\infty E[D^2|R = r]f_R(r)\,dr \\
&= \int_0^\infty \frac{r^3\pi\lambda(\mathbf{r})}{8}\exp[-\frac{\lambda(\mathbf{r})\pi r^2}{8}]\,dr = \frac{4}{\pi\lambda(\mathbf{r})}.
\end{aligned}
$$

**Cost Function:** Having determined the performance of the forwarding rule, we are now ready to specify the cost function for this microscopic model. For this, let us assume that a packet following the forwarding rule gets transported across an incremental distance of $\epsilon$. By the massively dense assumption, this distance is covered with a large number of hops. To calculate the associated incremental cost, we note that the sequence of progresses $\{X_i\}$ together with their associated costs $\{C_i\}$ form a reward renewal process, where $C_i$ is the reward of taking a step of size $X_i$. By the strong law of large numbers for reward renewal processes [15], it follows that $\frac{\sum_{i=1}^n C_i}{x} \to \frac{E[D^2]}{E[X]}$ as $x, n \to \infty$. By the massively dense assumption, the limit is achieved for $x = \epsilon$, therefore the incremental cost $\sum_{i=1}^n C_i = \frac{E[C]}{E[X]}\epsilon = \frac{4\sqrt{2}}{\pi}\frac{1}{\sqrt{\lambda(\mathbf{r})}}\epsilon$. To keep the expressions simple, let us assume that $c_{\text{nominal}} = \frac{4\sqrt{2}}{\pi}$. It follows that

$$c(\mathbf{r}) = \frac{1}{\sqrt{\lambda(\mathbf{r})}}.$$

The result is intuitive. Indeed, let us assume that the density is increased by a factor of 4. Then the expected length of a transmission decreases by a factor of 2, hence the area occupied by a transmission decreases by a factor of 4. However, now we need twice as many transmissions to cover the same distance. Therefore, if we increase the node density by a factor of 4, the overall area consumed decreases by a factor of 2. This back-of-the-envelope calculation is consistent with the cost function we arrived at.

### B. Energy Limited Networks

Let us now consider a network of nodes communicating over a common wireless channel with very large, practically
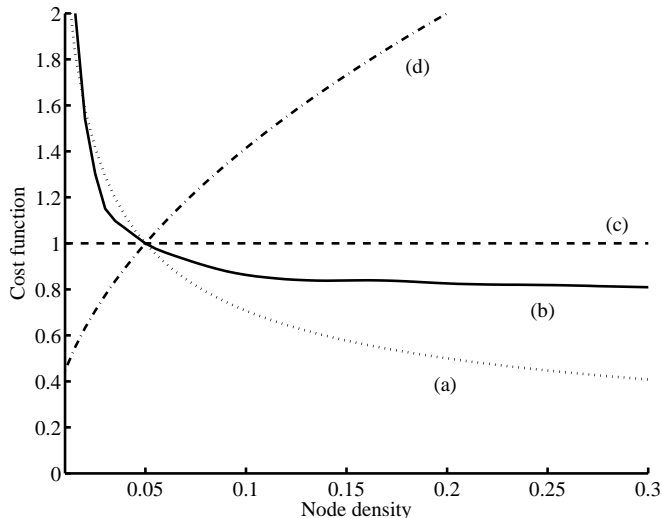
Fig. 6. The cost function versus the density for four microscopic models: **(a)** Bandwidth Limited, **(b)** Energy Limited, **(c)** Nominal network, **(d)** Minimum hop model of [5]. Costs are normalized so that they all intersect at the point $(0.05, 1)$.

infinite bandwidth. This is the case, for example, with nodes using Ultra-Wide-Band transceivers. In such a case, each transmitter can use its own dedicated bandwidth, and communication is not hampered by interference [20].

On the other hand, we assume that nodes have a limited energy supply. Therefore, we would like from the network to transport data in an energy efficient manner. We assume that nodes communicate using a fixed data rate, and in order to reach a receiver at a distance $d$, the transmission incurs an energy cost

$$C(d) = ad^b + c. \tag{7}$$

This is the standard model used in the literature [21], [22]. The constant term $c$ captures the energy consumed in the electronics of the transmitter and receiver, which is independent of the transmission distance. The term $ad^b$ captures the energy consumed by the transmitter power amplifier, and which depends on the distance. The exponent $b$ typically obtains its value in the range $(2, 6)$, depending on the environment [23].

Under this energy consumption model, transporting a packet across the network using a large number of very short hops incurs a very large aggregate energy cost, due to the constant term $c$. Likewise, transporting a packet with a few long hops also incurs a large energy cost, due to the term $ad^b$ which typically increases with distance much faster than linearly. In fact, it is intuitively clear that if the nodes could decide on the hop distance, they would select the optimal value $d_{\text{opt}}$ for which the ratio $\frac{ad^b+c}{d}$ is minimized. Straightforward calculation shows that $d_{\text{opt}} = \left[\frac{c}{(b-1)a}\right]^{\frac{1}{b}}$.

The forwarding Rule B used in the bandwidth limited model continues to make sense in this case, with the modification that the transmission cost is given by (7). As in the previous case, we would like to determine the expected cost and progress per

hop incured by this forwarding rule, and use them to establish the cost function.

A complication arises from the fact that the constant cost-progress loci $\frac{ad^b+c}{x} = r$ are no longer circles. In Fig. 4 we plot the loci for a typical choice of parameter set. As shown in the figure, the loci are oval shaped, roughly centered at the point $(d_{\text{opt}}, 0)$. As a result, the method of the previous section for finding the expected cost and progress can not be replicated. Nevertheless, we can easily calculate the expected cost and progress numerically, by Monte Carlo simulation. In Fig.6 we plot the resulting cost function versus the node density.

As in the previous case, the resulting cost function is a decreasing function of the node density. Contrary to the previous case, however, as the node density goes to infinity, the cost function converges to a strictly positive limit. This is expected: After the node density achieves a sufficiently high value, each node is guaranteed to find a node further down the trajectory and at a distance very close at the optimal $d_{\text{opt}}$. After that point, increasing the density further does not affect the cost function, as the optimal point of operation has already been achieved.

*C. Other Models*

Our bandwidth limited and energy limited models are only two of the many possibilities. Using the same methodology, many other choices are possible. We briefly mention that in [5] the cost function $c(\mathbf{r}) = \sqrt{\lambda(\mathbf{r})}$ was implicitly used. As discussed, that cost function corresponds to the case where we want to minimize the number of hops, while restraining the communication between nearest neighbors. An alternative cost function was also introduced in [11], under a formulation less developed that our own. That cost corresponded in the case where we want to minimize the energy needed for the transport of a packet, but all nodes are required to transmit with fixed power, i.e. the parameter $a$ of (7) is set to $a = 0$. The authors of [24] propose, independently from our line of work, a cost function that captures the energy needed for routing in a large-scale static interference field.

Alternatively, the cost function could be specified in an adaptive manner, depending on the fluctuating needs and capabilities of the network. For example, we could perform adaptive routing by using a cost function that captures the congestion at the various parts of the network. Therefore, congested areas could select a high value for their cost function, and so discourage the use of routes through them. Areas of the network with spare resources could select low values for their cost function, thereby inviting routes through them. This is a potential subject for future work.

V. SIMULATION RESULTS

In this section, we present a few examples of optimal routes, under the different microscopic models we have considered. Due to space constraints, this section is only intended to provide a proof-of-concept and illuminate the theory, rather than exhaustively investigate the performance of protocols based on the Optics analogy.
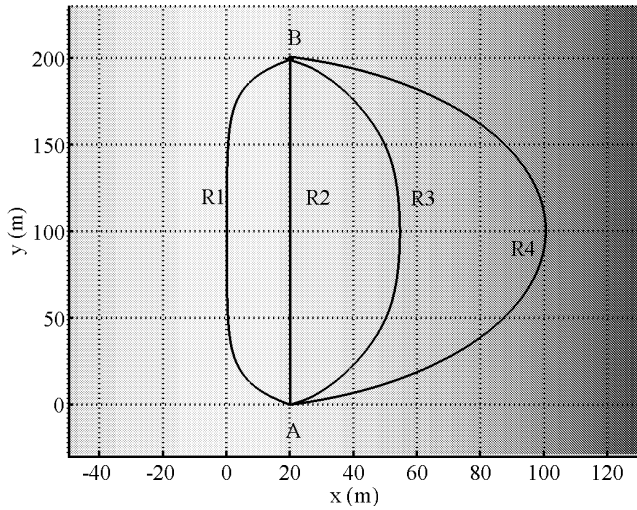
Fig. 7. Optimal routes under four microscopic network models: **(R1)** The model of [5], **(R2)** Constant cost function model **(R3)** Energy limited model, **(R4)** Bandwidth limited model.

We consider a network in which the node density $\lambda(\mathbf{r}) = 3 \times 10^{-5}x^2 + 0.01$, i.e., there is no dependence of the density on the $y$-coordinate. In Fig. 7 we plot the optimal routes connecting the two points $A = (20, 0)$ and $B = (20, 200)$ for the bandwidth-limited network model (Route R4), the energy limited model (route R3), the model of [5] (Route R1), and a network in which the cost function is constant and independent of the node density. In the last case, the optimal route is simply the straight line connecting the points $A$ and $B$. In the figure, areas with high node density appear darker.

As expected, under the cost model of [5], the optimal route goes through the region of the network with very low density. Also as expected, in both the energy-limited and the bandwidth-limited cases the optimal routes (R4 and R3 respectively) are through the part of the network with high concentration. However, R3 does not deviate as much as R4. The reason is that, in the bandwidth limited case, the cost function $c$ decreases with the density for arbitrarily high values of the density $\lambda$ as $c = \frac{1}{\sqrt{\lambda}}$ (see Fig. 6). For the energy limited case, however, after the density reaches some value, a further increase does not reduce the cost. Therefore, deviating through regions of very high density is not beneficial and the optimal routes change accordingly.

For the bandwidth limited case, the aggregate cost of the straight line R2 is $80\%$ more than the aggregate cost of the optimal route R4, even though R4 is significantly longer. Similar gains exist with respect to the straight line for the other cases.

In Fig. 8, we plot wavefronts and rays in an example of distributed routing and for two microscopic models: the bandwidth limited, and the model of [5]. The region $\Omega$ is a circle centered at the point $(20, 0)$, with radius 10. As expected, under the bandwidth limited model, regions of low node density incur high costs, and for this reason routes bend so that they pass this region as fast as possible. The opposite
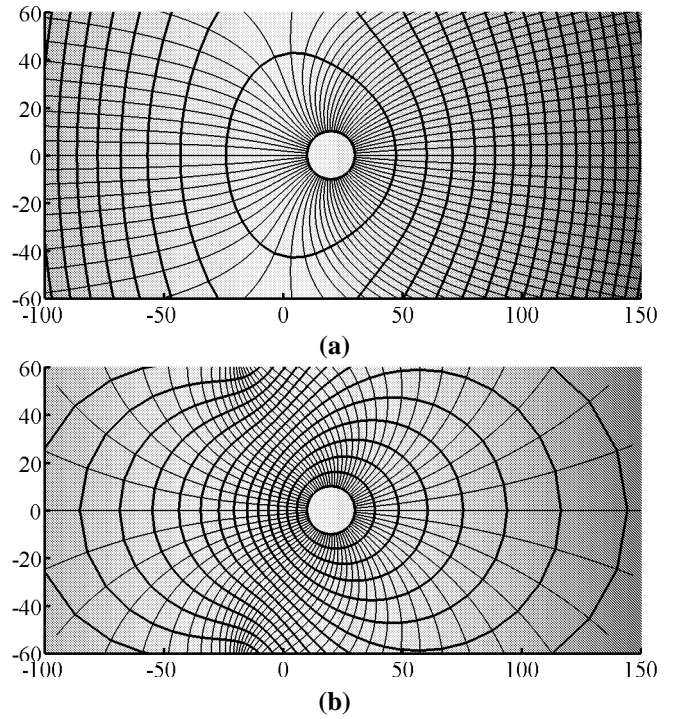


Fig. 8. Wavefronts and rays in a case of distributed routing, and for three microscopic network models: **(a)** The model of [5], **(b)** Bandwidth limited model.

happens for the model of [5].

Finally, in Fig. 9 we consider a particular realization of a network with node density $\lambda(\mathbf{r}) = 0.5 \times 10^{-4}x^2 + 0.025$, and we compare *(i)* the route that results from performing TBF according to Rule B on the route determined by the Optics analogy, with *(ii)* the route that results from performing TBF on the straight line connecting the source and the destination, and *(iii)* the optimal route, determined by the Bellman-Ford algorithm. We note that the optimal route follows closely the route determined by Optics. That should be expected, as our route provably becomes optimal in the limit of a large number of nodes. The costs of the three routes are 1305, 2750, and 944 respectively. Therefore, optical routing needs less than half the resources needed when routing along the shortest path, but $40\%$ more resources with respect with the optimal route. We note, however, that calculating the optimal route becomes more and more impractical as the number of nodes increases.

## VI. CONCLUSIONS

In this work, we investigate the problem of optimally selecting routes in massively dense networks. We use the observation of [5] that the optimal routing problem can be cast as a problem in Optics, significantly extending the theory and methodology of that work. Although in practice it may not be possible to implement our routing algorithms in networks with modest size and computing capabilities, still, knowing the limiting optimal routes allows a meaningful evaluation of simpler, suboptimal algorithms, and provides useful insights for the development of appropriate routing rules.
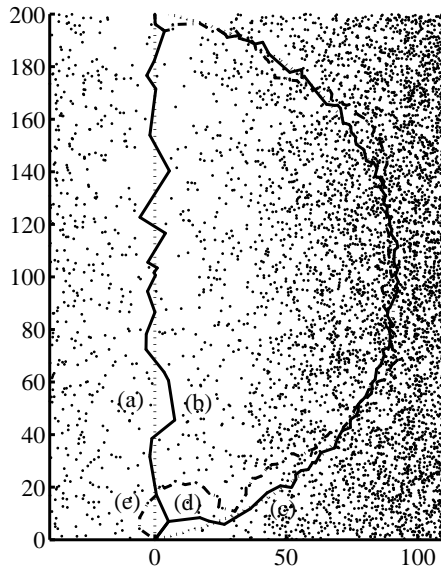
Fig. 9. A realization of a massively dense network, with roughly 5000 nodes, and with a source placed at $(0,0)$ and a destination placed at $(0,200)$. **(a)** Straight line connecting the source and destination. **(b)** Trajectory based forwarding along the straight line. **(c)** Curve determined by the Optics analogy **(d)** Trajectory based forwarding on curve (c). **(e)** Optimal route determined by Bellman-Ford algorithm.

Due to lack of space, a number of important issues remain unaddressed. First of all, although Geometrical Optics describes optimal routes in a very elegant manner, it does not provide the tools for a source-destination pair to discover, in a distributed manner, the optimal route connecting them. The situation is complicated by the fact that it is possible, in heavily inhomogeneous media, to have multiple rays connecting two points. The principle of Fermat continues to hold, and all rays will be *locally* optimal, but not necessarily *globally* optimal. Mathematical speaking, this implies that the eikonal function has multiple branches. In the context of networking, this means that it is possible to have multiple routes connecting a source-destination pair, each of which achieved a local minimum of cost, but not necessarily the global minimum. These practical issues are addressed in [25].

REFERENCES

[1] C. Perkins, Ed., *Ad Hoc Networking*, 1st ed. Boston, MA: Addison-Wesley, Dec. 2000.
[2] C. Santinavez, B. McDonald, I. Stavrakakis, and R. Ramanathan, "On the scalability of ad hoc routing protocols," in *Proc. IEEE INFOCOM*, vol. 3, New York, NY, June 2002, pp. 1688–1697.
[3] D. Niculescu and B. Nath, "Trajectory based forwarding and its applications," in *Proc. ACM MobiCom*, San Diego, CA, Sep. 2003, pp. 260–272.
[4] S. Giordano, I. Stojmenovic, and L. Blazevic, "Position based routing algorithms for ad hoc networks: A taxonomy," *Ad Hoc Wireless Networking*, pp. 103–136, 2004.
[5] P. Jacquet, "Geometry of information propagation in massively dense ad hoc networks," in *Proc. ACM MobiHOC*, Roppongi Hills, Japan, May 2004, pp. 157–162.
[6] M. Kalantari and M. Shayman, "Energy efficient routing in wireless sensor networks," in *Proc. Conference on Information Sciences and Systems*, Princeton University, NJ, Mar. 2004.
[7] S. Toumpis and L. Tassiulas, "Packetostatics: Deployment of massively dense sensor networks as an electrostatics problem," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005.
[8] ——, "Optimal deployment of large wireless sensor networks," *IEEE Trans. Inform. Theory*, Jul. 2006.
[9] B. Sirkeci-Mergen and A. Scalione, "A continuum approach to dense wireless networks with cooperation," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005.
[10] E. Hyytiä and J. Virtamo, "On load balancing in a dense wireless multihop network," in *Proc. 2nd EuroNGI Conference on Next Generation Internet Design and Engineering (NGI 2006)*, València, Spain, Apr. 2006.
[11] R. Catanuto and G. Morabito, "Optimal routing in dense wireless multihop networks as a geometrical optics solution to the problem of variations," in *Proc. IEEE ICC*, Istanbul, Turkey, June 2006.
[12] M. Born and E. Wolf, *Principles of Optics*, 7th ed. New York, NY: W. H. Freeman and Company, 1999.
[13] E. Hecht and A. Zajac, *Optics*, 4th ed. Addison Wesley, 2003.
[14] J. A. Sethian, *Level Set Methods and Fast Marching Methods*, 1st ed. Cambridge University Press, 1999.
[15] S. Ross, *Stochastic Processes*. New York, NY: John Wiley and Sons, Jan. 1995.
[16] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the gap in the capacity of random wireless networks," in *Proc. IEEE ISIT*, Chicago, IL, June-July 2004.
[17] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.
[18] S. Lee, B. Bhattacharjee, and S. Banerjee, "Efficient geographic routing in multihop wireless networks," in *Proc. ACM MobiHOC*, Urbana-Champaign, IL, May 2005, pp. 230–240.
[19] L. Kleinrock and J. Silvester, "Spatial reuse in multihop packet radio networks," *Proc. IEEE*, vol. 75, no. 1, pp. 156–167, Jan. 1987.
[20] R. Negi and A. Rajeswaran, "Capacity of power constrained ad-hoc networks," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar.-Apr. 2004.
[21] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM*, vol. 1, Tel Aviv, Israel, Mar. 2000, pp. 22–31.
[22] A. Ephremides, "Energy concerns in wireless networks," *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, pp. 48–59, Aug. 2002.
[23] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 1st ed. Upper Saddle River, NJ: Prentice Hall, 1996.
[24] F. Baccelli, N. Bambos, and C. Chan, "Optimal power, throughput and routing for wireless link arrays," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.
[25] R. Catanuto, G. Morabito, and S. Toumpis, "Optical routing in massively dense networks: Practical issues and dynamic programming interpretation," in *Proc. International Symposium on Wireless Communications Systems*, Valencia, Spain, Sep. 2006.